

Article

Not peer-reviewed version

Profit From Time Decay In Options Using Deep Learning

[Sumit Kumar](#)*, Andrés Quintero*, Joe Wayne Byers

Posted Date: 24 December 2024

doi: 10.20944/preprints202412.1934.v1

Keywords: Options; option price; time value; theta decay; deep learning; reinforcement learning; machine learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Profit From Time Decay In Options Using Deep Learning

Sumit Kumar, Andrés Quintero and Joe Wayne Byers

WorldQuant LLC, WorldQuant University; pahwa.sumit@gmail.com; andres.quintero.wqu@gmail.com; joe.byers@wqu.edu

Abstract: One of the most flexible assets in the financial markets is the options, the flexibility makes it possible to create virtually unlimited strategies among different underlying assets, and combinations, contrary to the stock market, where the mechanism to sell or buy is relatively simple, other elements, like time can play against the trader. The derivative market could be overwhelming considering the broad alternatives, with high entry barriers and a steep learning curve. The new traders try to jump these barriers without a plan, hurrying to place trades that follow their emotions, and losing money along the process, as the best alternative to learn from trial and error. This research proposes an options trading framework that can help reduce this gap and profit from time decay.

Keywords: Options; option price; time value; theta decay; deep learning; reinforcement learning; machine learning

1. Introduction

The premium of an equity option is the sum of its intrinsic value, based on the difference in the underlying price, and its time value, which decreases as expiration approaches (time decay), and it can benefit or harm the option holder. Some strategies use time to their advantage, like Credit Spreads and Short Straddle.

This work aims to reduce the barrier entries for new traders using a quantitative approach, this research proposes to develop an options trading framework for profiting from time decay and using deep learning algorithms. The system will be able to identify the best trading strategies with the help of a hybrid deep learning model that uses Long-short Term Memory together with Convolutional Neural Networks and will be used to forecast the underlying asset price trend, and profit from theta decay. It will select the entry and exit criteria for maximum profit by analyzing the data and using reinforcement learning. The project will be aimed at data collection, preparation, data analysis and feature selection, model tuning, optimization, and performance evaluation using a backtest approach.

This approach attempts to systematically implement options strategies with the help of deep learning, which requires extensive analysis and trading experience to make decisions.

2. Literature Review

The premium is the value of an equity option, and it is the sum of intrinsic value plus the time value, the first one depends on the difference in the underlying price, and the second is directly related to the time to expiration as defined by Morris's guideline [11]; the more time the option to be in-the-money, the more time value. The drop in value is not linear, as it approaches expiration, it drops quickly (the generally accepted idea). This decline in value is also called "time decay", and the option status can play in favor or against you as an option holder. Theta is an option Greek that measures the rate at which a premium decays over a unit of time according to Morris [11].

Many Authors give insights and rules to profit from the theta decay for pointing out strategies [13], without a method to trade systematically, and most of these approaches are based on experience and sentiment with no quantitative support. Words like fearful, think, and expect, appeared in those

articles [14]. Some strategies use time to their advantage, like Credit Spreads, which rely on the directional movement of the stock, and theta decay would be in favor defined in Cohen's book [2]. The Short Straddle and Strangle (a modification to the Straddle) are both directional neutral, benefit from low volatility, and time decay is an advantage because it is a short position. The risk is uncapped as explained by Cohen [2].

One interesting finding about time decay patterns is that, contrary to the spreading wisdom that says time decays slowly at the beginning and speeds up close to the expiry day, an empirical study carried out by McKeon [10], finds that there is a consistent pattern of time decay between calls and puts with similar moneyness, the ATM options have significant time decay early, regardless of expiry date, and for IT and OTM options, time decay is slower but shows a sharp decline on the final day. This study enforces the need for new ways to discover the complex dynamics of the option price and time decay, that do not rely only on BSM (Black Scholes Model) without considering how the underlying asset price impacts intrinsic value and time value.

The application of machine learning is not the exception in options, and one of the uses is pricing, where the current numerical methods are slow or there is a lack of resources and used by Li [8]. The author also points to Long-short Term Memory (LSTM) as one of the most used methods for learning time series, and the combination of traditional methods like Markov Decision Process (MDP) to find the strategy and using Artificial Neural Networks (ANN) with Support Vector Machine (SVM) to know how to trade, this work has limited application, and more verifications will be needed. Another application is the improvement of accuracy and efficiency using Physics-Informed Neural Networks (PINN) compared with Partial Differential Equations (PDE) to solve BS used by Wang, Li and Li [15] in option pricing.

Deep learning applications and option portfolios are not the exception in this field (Xu and Ma) [18], by using a model that sells contracts and predicts future underlying asset prices, the model can find optimal portfolio composition. Additionally, this work considers risk management, which serves as a reference for the wide scope of ML/DL uses.

Another application has developed a trading system that uses the Kelly criterion with machine learning methods (SVM, Random Forest, KNN, and Naive Bayes), and ensemble the win rate results, Wu and Chen [17] show that it can generate profits while controlling the risk, this kind of application can serve as inspiration for our work. The decision tree classifiers have been used by Joshi, Venkateswaran, and Bhattacharyya [6] to classify the best strategy where technical indicators data can be used along with options Greeks, and the implied volatility to find the best strangle strategies from possible combinations. Some advanced methods also suggest using the combination of reinforcement learning with deep learning for better results as per Hilpisch [5]. Researchers Wen, Yuang and Yang [16] have also used reinforcement learning in options trading, and propose a framework that address the problem of insufficient data, that is common issue in finance applications, and also, they used three actions: "Do nothing", "Buy an asset", and "Close the position", future work proposed multi-leg strategies that are not covered in their work.

The underlying asset price is fundamental in options trading, the profit at maturity equation will involve this value. Researchers have used hybrid deep learning models like Convolutional Neural Networks (CNN), LSTM, and Gated Recurrent Units (GRU) to stock forecasting, these works then have relevance also for options trading, and hybrids models have shown to be effective as per Naufal and Wibowo [12].

The options Greeks provide (purely mathematical) insights about the strategy selection. Adding technical analysis of the underlying stock gives more accurate ways to decide the strategies according to Kakushadze and Serur [7]. The reason is that any option contract is a derivative of an underlying stock, and the latter price movement prediction is critical to determining any option strategy.

Fundamental analysis of the underlying stock helps to find intrinsic value by looking at all financial aspects of that company. It assumes that the stock market might not reflect the actual price of the underlying stock, in the long run, it will reflect the price fundamentals as Drakopoulou points

out [3] and revert to the mean. Screening an underlying stock based on fundamentals can attract more participants, which will further induce more liquidity in the market.

Technical analysis uses historical prices, volume, chart patterns, and leading/lagging indicators as it relies on some assumptions like: “the market discounts everything”, “stock price moves on trends”, and “stock price trend history repeats itself” made by Drakopoulou [3]. One can figure out the stock price (bullish, bearish, or neutral) direction by using technical indicators according to Kakushadze and Serur [7]. Time is a precious entity while trading options, a lagging indicator will cause a delay, and it will miss the right entry or exit points while executing the strategy. The recommended leading indicators by Ciana [1] are as follows:

- Relative Strength Index (RSI).
- Moving Average Convergence Divergence (MACD).
- Bollinger Bands (BB).
- Directional Movement Index (DI+, DI- and ADX).

It will be possible to classify the regimes based on Majidi, Shamsi, and Marvasti [9] using technical indicators, which would help to decide the direction of the underlying price.

The risk numbers are reflected in the implied volatility and Greeks i.e. delta (sensitivity in option prices due to change in underlying price), theta (sensitivity to time expiry), gamma (sensitivity to change in Delta), vega (sensitivity to IV) and rho as per Cohen [2]. It will leverage the regime switch of underlying, delta, theta, and implied volatility from the available option chains to find the best strategy with low risk and high profit, i.e. a better Sharpe ratio.

Our approach attempts to systematically implement options strategies that, without the help of quantitative analysis and deep learning tools, require long trading experience to make decisions.

2. Theoretical Framework

2.1. Options Strategies

Options trading strategies are an effective tool for managing various assets; it is granted to stock traders in case of generating subsequent profits upon the analysis of the stock market fluctuations and theta decay. The Credit Spread is one such strategy with 2 legs. It consists of selling one OTM option and buying another OTM option with the same expiration date but a different strike price (the gap is called the spread). It offers both defined risk and profit potential.

Credit spread is of two types: The Bull Put Spread (BPS) shown in Figure 1 and Bear Call Spread (BCS) depicted in Figure 2. More specifically, the bull put spread involves executing a long Out of the money Put as well as a short Out of the money Put having a higher strike price. This is a bullish trade, and the trader will receive credit for carrying it out. If one is expecting the underlying movement to be neutral, then one can have a Short Strangle as Figure 3 where they sell the OTM call and put it together.

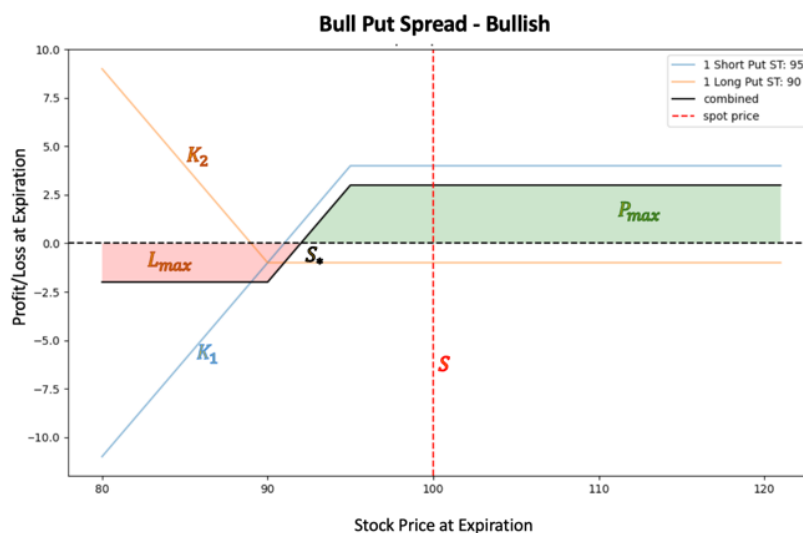


Figure 1. Options trading strategy Bull Call Spread - Bullish.

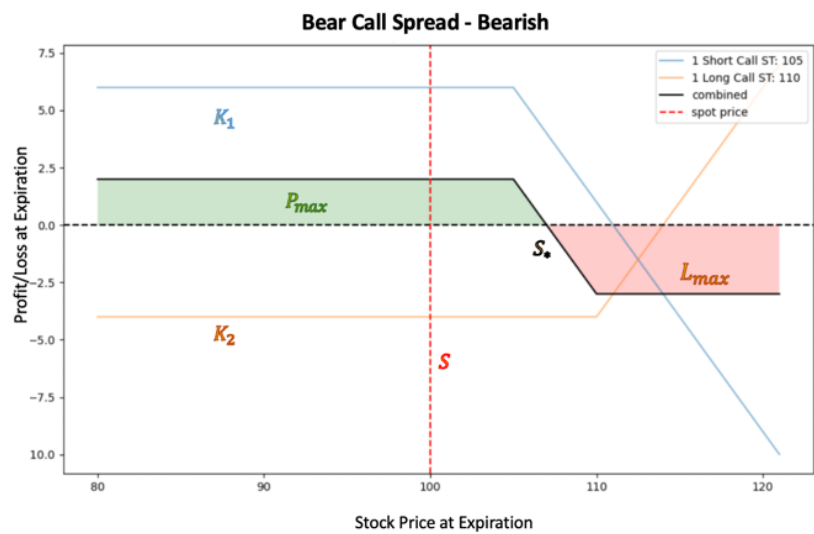


Figure 2. Options trading strategy Bear Call Spread - Bearish.

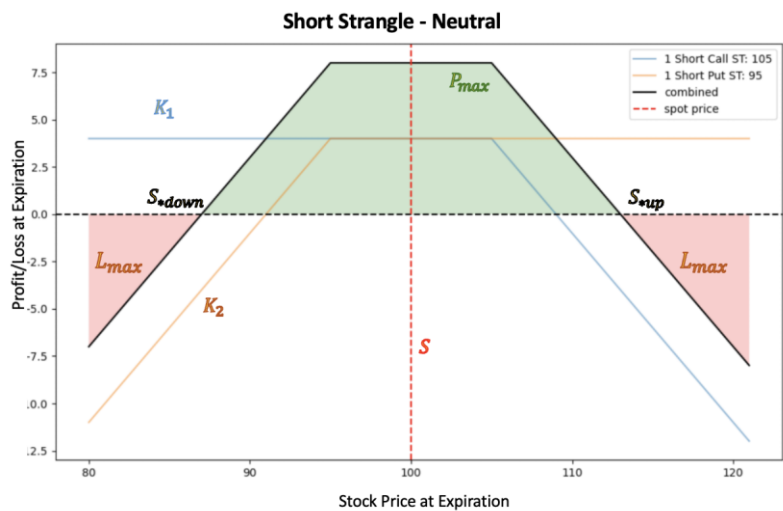


Figure 3. Options trading strategy Short Strangle - Neutral.

As per Kakushadze and Serur [7], the following Table 1, explains the maximum profit, maximum loss, and break-even for each strategy.

Table 1. Option strategy credit spread.

Strategy	Equations	Definitions
Bull Put Spread	$F_T = (K_1 - S_T)^+ - (K_2 - S_T)^+ + C$	(1) F_T is the payoff at maturity T
	$S_* = K_2 - C$	(2) S_T is the stock price at maturity T C is the net credit received at $t=0$
	$P_{max} = C$	(3) S_* is the break-even price
	$L_{max} = K_2 - K_1 - C$	(4) P_{max} is the maximum profit at maturity L_{max} is the maximum loss at maturity

Bear Call Spread	$F_T = (S_T - K_1)^+ - (S_T - K_2)^+ + C$	(5)	F_T is the payoff at maturity T
	$S_* = K_2 + C$	(6)	S_T is the stock price at maturity T C is the net credit received at $t=0$
	$P_{max} = C$	(7)	S_* is the break-even price
	$L_{max} = K_1 - K_2 - C$	(8)	P_{max} is the maximum profit at maturity L_{max} is the maximum loss at maturity
Strangle	$F_T = -(S_T - K_1)^+ - (K_2 - S_T)^+ + C$	(9)	F_T is the payoff at maturity T
	$S_{up} = K_1 + C$	(10)	S_T is the stock price at maturity T C is the net credit received at $t=0$
	$S_{down} = K_2 - C$	(11)	S_{up} is higher break-even
	$P_{max} = C$	(12)	S_{down} is lower break-even P_{max} is the maximum profit at maturity
	$L_{max} = Unlimited$	(13)	L_{max} is the maximum loss at maturity

Trade with options requires a deep understanding of time decay impact (theta decay), and it is a crucial factor in making profits in the financial derivative market. Some strategies can take this advantage and use the time in favor. Traditional methods rely on BSM, new studies suggest that the use of ML/DL methods can improve not only the precision but also the needed computational resources. There is a benefit to using systematic methods and machine learning in options markets, considering the wide alternatives to trade, understanding the hidden patterns in time decay, as well as implementing complex strategies that are possible only by experienced traders, opening an application chance in a competitive field. The combination of well-known technical indicators, with fundamental analysis, makes it possible to analyze more information than a human can do, removing psychological and sentiment factors that take decades to enhance.

2.2. LSTM-CNN

LSTM is a kind of Recurrent Neural Network (RNN) that uses cell memories and is not prone to “vanishing gradient problems” that do not let the network learn long-term dependencies; this characteristic makes it suitable for time series forecasting. An LSTM cell based on Ravichandiran’s structure [4], it is depicted in Figure 4.

The decisions about how long the information is retained, it is taken based on the below parts:

- Forget gate: This is responsible for deciding the information that should not be in the cell state C_t .
- Input gate: This gate decides which information is stored in the memory.
- Output cell: It is responsible for the information that is shown from the cell state C_t at time t .

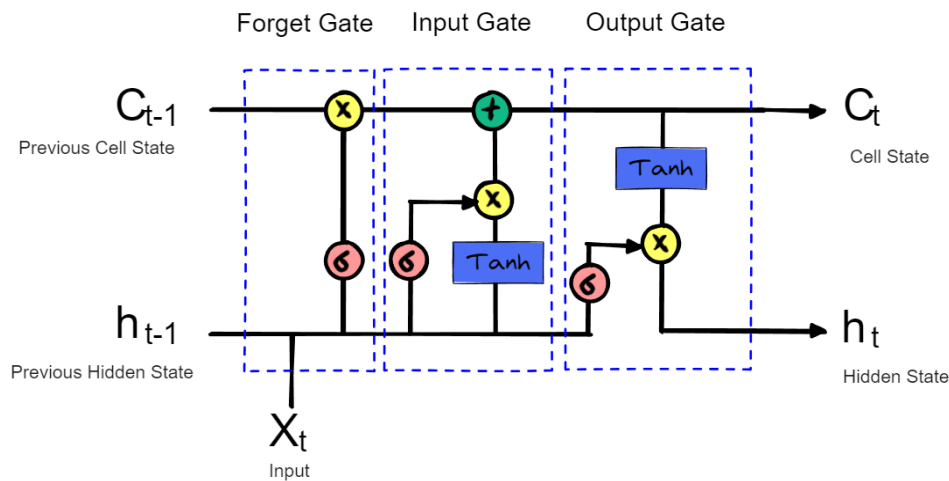


Figure 4. LSTM Cell.

CNN is a class of Neural Networks that is highly used in areas like computer vision, and it has been efficiently used also for time series to capture temporal patterns. The main elements of CNN are the Convolutional layer, Pooling layer, and fully connected layer as in Figure 5.

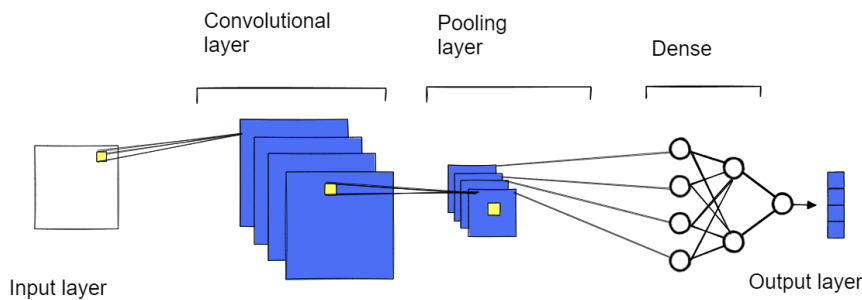


Figure 5. CNN layers.

Combining LSTM and CNN can capture long-term dependencies as well as extract patterns, and features from the data. Other advantages of the hybrid models are their improved robustness and precision.

2.3. Reinforcement Learning

It is a method where an agent learns from the environment interaction, with each action the system tries to increase the reward using the feedback received from previous actions and adjust its strategy to perform better as defined by Ravichandiran [4]. This machine learning method can be employed to make complex decisions in options trading like entry, hold, and exit criteria to avoid significant losses (risk management). The agent can learn from historical data, and adapt to get the highest reward, the observations are the current market conditions, underlying asset price, volatility, Greeks, and other indicators, the policy maps observations to actions, and looks for profit maximization, finally, the policy is the strategy to know which action should be taken (entry, hold and exit). The overall diagram can be seen in Figure 6.

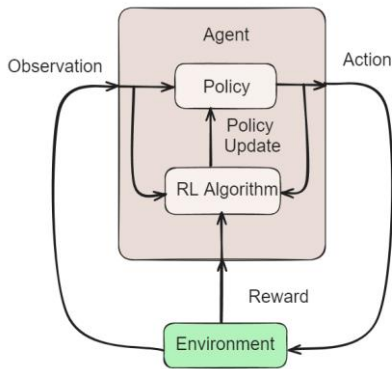


Figure 6. Reinforcement Learning diagram.

3. Methodology

This research proposes a framework with the following elements that aim to profit from theta decay and systematically trade on options. The main components and techniques are detailed in Figure 7 below.

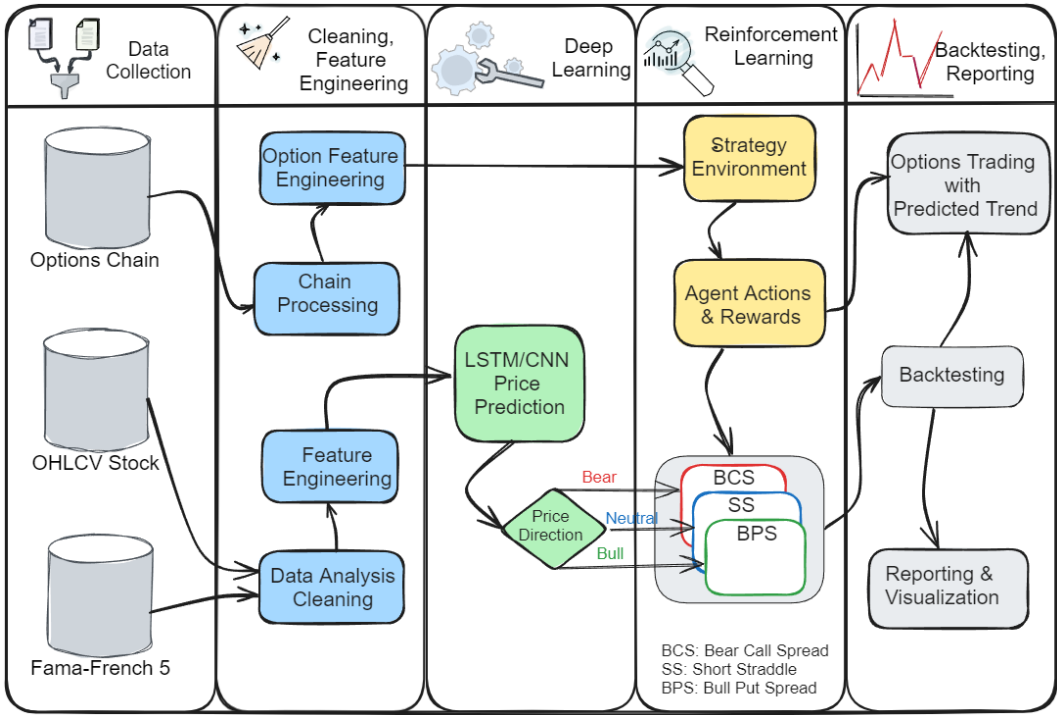


Figure 7. Options Trading Framework.

3.1. Data Collection

The data used in this research is the underlying asset data OHLCV (Open, High, Low, Close, Volume) and other relevant information for AAPL (Apple Inc) from the US Stock Market NASDAQ. This data, spanning from Aug. 2018 to Dec. 2023, will also be used to construct the technical indicators. The Fama-French 5 Factors data was also added for the same period, this can be considered as fundamental data.

Daily Option data for all available chains during that period is collected, and Table 3 summarizes the data.

Table 3. Data sources AAPL asset.

Data	From	To	Periodicity	Source
OHLCV	2018-08-27	2023-12-14	Daily	Yahoo Finance
Historical Options Chain	2020-01-09	2023-10-31	Daily (with monthly expiries)	OptionsDX
Future Options Chain	2023-01-11	2023-12-31	Daily (with monthly expiries)	Yahoo
Fama-French Factors	2018-08-27	2023-12-14	Daily	FF Research Data 5_Factors_2x3_daily

3.2. Feature Engineering

The data is cleaned by filling in missing information. Technical indicators are calculated for use in the price prediction. The following provide details to this step.

- Relative Strength Index (RSI): This momentum indicator will give us insights about the uptrend or downtrend in the price, the direction and strength in price movement. The default configuration is to use 14 periods. It has a range value between 0 and 100.
- Moving Average Convergence Divergence (MACD): This is another momentum indicator. It is calculated by subtracting 12 and 26 EMA periods on closing prices. It identifies strengths, directions, and momentum in stock prices.
- Bollinger Bands (BB higher, low): The BB are used to know if the prices are high and low in relation to each other. This indicator can measure volatility and also trends in the stock movement.
- Directional Movement Index (DI+, DI- and ADX): These indicators also measure strength and direction, it is used to confirm trends, and it can be used together with ADX (Average Directional Index) to show momentum.
- Exponential Moving Average (EMA): This indicator gives us information about price change direction, values from 30 and 50 days are considered.
- On-Balance Volume (OBV): OBV helps to confirm uptrend or downtrend as regards price increase or the opposite, basically measuring the buying and selling pressure.
- Accumulation/Distribution Line (ADL): This measures the cumulative flow of money in and out of a stock, it belongs also to the volume group indicators.
- Aroon Indicator: It identifies changes in price trends, it is composed of two lines and their interactions give information about the strength of uptrend and downtrend.
- Average True Range (ATR): This is primarily used to measure volatility or average price range over time.
- CBOE Volatility Index (VIX): This reflects the market's expectations about volatility over the next 30 days, the volatility can be seen as a measure of risk in the market.

The Fama French 5 Factors Data was collected also, and can be considered as fundamental data, these factors try to explain the variation in the stock returns as per Fama and French [19]:

- Market Risk (Mkt-RF): This is the market excess returns over the risk-free rate (market premium).
- Size Factor (SMB - Small Minus Big): The returns between small-cap and large-cap stocks (size premium).
- Value Factor (HML - High Minus Low): The return spread between value and growth stocks (value premium).
- Profitability Factor (RMW - Robust Minus Weak): The return between stocks of companies with strong profitability and those with weak profitability.
- Investment Factor (CMA - Conservative Minus Aggressive): The return difference between conservative stocks investment companies and those that invest aggressively.

3.3 Deep Learning

The objective for this step is to take the features previously calculated and use them as input in a Deep Learning model that will be used as a price predictor. The main point of interest is a price range that will help with predictions weeks ahead, where the market could have a bull, bear, or neutral direction, then the corresponding option will be selected, Bull Put Spread, Bear Call Spread, or Short Strangle. Other tasks include input transformation, data split train/test/validation, model creation, hyperparameter tuning, sequence creation, and prediction. The proposed model is a combination of LSTM and CNN, the first one can capture long-term dependencies in the time series, and the second can capture patterns and features from the data, adding dropout layers helps for generalization, and finally, a concatenation layer to connect both legs, allowing the model to learn from both.

3.3.1 LSTM/CNN Model: The proposed LSTM/CNN is described as below, the number of components in the architecture and the final number of units in it is adjustable via hyperparameters tuning for the best fit, see the next Figure 8. The overall model architecture will have the following elements:

- Inputs: Stock data OHLCV (5), VIX, Fama-French (5) factors, and technical indicators (15).
- LSTM Leg (OHLCV): Two LSTM layers with 64 units each and a Dropout layer 0.2 to avoid overfitting.
- CNN Leg (technical indicators): It is composed of Conv1D layer 32 filters and a kernel size of 3, MaxPooling1D layer with a pool size of 2, Flatten layer to convert the 2D output to 1D, and Dropout with a rate of 0.2.
- Legs concatenation and Output: These previous two legs are concatenated, and connected with a Dense layer of 128 units, then Dropout 0.2, and finally output Dense layer with 20 (days to be forecasted).

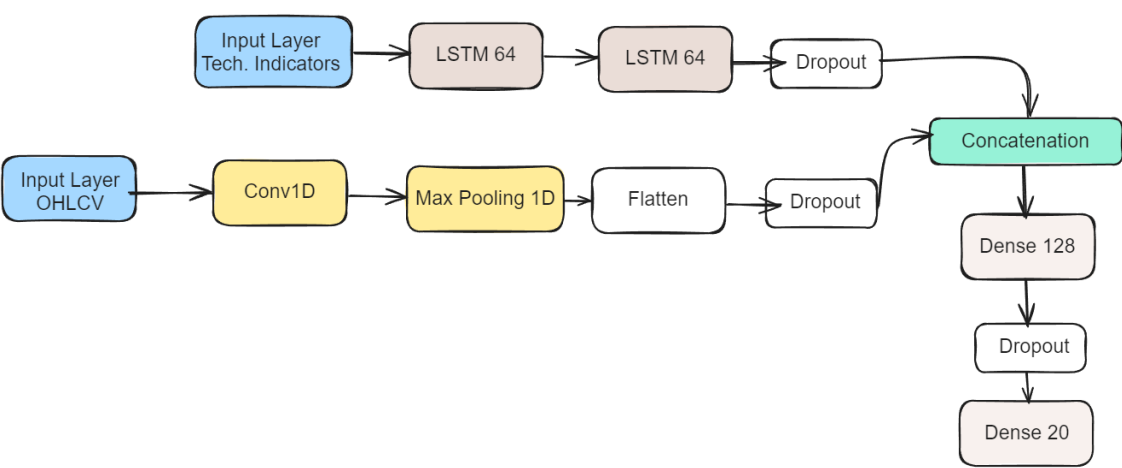


Figure 8. LSTM/CNN Network architecture for price prediction

3.4. Historical Options Chain Analysis

The US market provides both weekly and monthly expiration services for different financial instruments. Monthly expirations are the specific of this research. This approach will help us to get more meaningful results. OTM calls and puts from the superset of the monthly expiry options to profit from the theta decay. The Strategy Engine for this research has the following responsibilities:

- Analyze the historical options chains for AAPL and filter out the weekly expiry chains.
- Run exploratory data analysis to check and fix: The null values records, column names, data types of the columns, and zero volume records.

- For a given strategy, prepare the spreads/legs within the same expiry date and quote date for each combination of the lower and higher strike with the DTE, Greeks, IV, etc.

3.5. Feature Engineering for the Historical Options Chains

The following Table 4 consists of target columns, expected return on the paid margin (EXPECTED_RETURN_ON_MARGIN), and the Probability of expiring worthless (PROB_OF_EXPIRING_WORTHLESS), and the rest of the columns as the features.

Some of the key features are as follows:

- Net Credit: It is the difference between the bid price and the asking price of the legs.
- Max Profit: In the case of selling strategies, the maximum profit is always the net credit received when the trader enters the trade.
- Max Loss: The maximum loss or strategy margin is the difference between both strikes and credit received.
- Net Delta: It is the difference between the delta values of both legs.
- Probability of expiring worthless; As delta also signifies the probability of an option strategy expiring ITM (In the Money), the inverse probability of Net Delta tells the probability of expiring worthless.
- Margin paid: This is the margin paid to the broker, since the strategies are short, so margin requirements would be higher and could be different in each strategy type.
- Expected Return on Margin: This is the expected return for the paid margin, which is basically the max profit divided by the paid margin.

Table 4. Options Chains’ Feature engineering.

Bull Put Spread	Bear Call Spread	Short Strangle
HIGHER_STRIKE	HIGHER_STRIKE	HIGHER_STRIKE
LOWER_STRIKE	LOWER_STRIKE	LOWER_STRIKE
UNDERLYING_LAST	UNDERLYING_LAST	UNDERLYING_LAST
QUOTE_DATE	QUOTE_DATE	QUOTE_DATE
EXPIRY_DATE	EXPIRY_DATE	EXPIRY_DATE
SHORT_PUT_BID	SHORT_CALL_BID	SHORT_PUT_BID
SHORT_PUT_DELTA	SHORT_CALL_DELTA	SHORT_PUT_DELTA
LONG_PUT_ASK	LONG_CALL_ASK	SHORT_CALL_ASK
LONG_PUT_DELTA	LONG_CALL_DELTA	SHORT_CALL_DELTA
MAX_PROFIT	MAX_PROFIT	MAX_PROFIT
NET_DELTA	NET_DELTA	NET_DELTA
PROB_OF_EXPIRING_WOR	PROB_OF_EXPIRING_WOR	PROB_OF_EXPIRING_WOR
THLESS	THLESS	THLESS
DTE	DTE	DTE
MAX_MARGIN	MAX_MARGIN	MAX_MARGIN
EXPECTED_RETURN_ON_	EXPECTED_RETURN_ON_	EXPECTED_RETURN_ON_
MARGIN	MARGIN	MARGIN

3.6. Strategy Execution

Reinforcement Learning is used to trade the given strategy. A custom environment is created for each strategy type (spreads/strangle) to simulate trading scenarios. It trains the model using the Proximal Policy Optimization (PPO) algorithm with Multilayer Perceptron Policy (MlpPolicy), a feedforward artificial neural network. The environment feeds back into the options chain historical information to improve as well as analyze the trading strategies.

3.6.1. Environment Setup

The custom environment based on OpenAI Gym, Gymnasium, and its extension Stable Baseline3. It includes the columns in the data frame as the observations, and two discrete actions (open and avoid positions). The reward is calculated based on predefined criteria related to days to expiration, implied volatility, return on investment, and the probability of options expiring worthless. Actions and corresponding rewards are as follows:

- Action 1 (Open): Rewards positively high probability if expiring worthless and substantial expected return on the paid margin and days to expiration between 15 and 45.
- Action 0 (Avoid): Else reward negative for the rest of combinations.

4. Results

4.1. Underlying asset trend prediction

AAPL stock data is cleaned, and Fama-French Factors, the resulting N/A values from technical analysis (EMA200) were removed, and the effective date range was between 2019-06-13 and 2023-05-19. This data was used for training and validation of the hybrid LSTM-CNN model, see Table 5, the data is divided into two groups, the first group OHLCV, EMA200 feeds the LSTM leg, and the remaining data feeds the CNN leg.

Hyperparameter tuning was performed in the deep learning model and using a search space as the Table 6. The best model was trained, and the learning curve (Appendix Figure 13) shows that training loss has decreased significantly over time, while the validation loss seems to be flattening, it could be due to the underfitting of the data.

Table 5. Data split.

Data	From	To	Model
Training/Validation	2018-08-27	2023-05-22	LSTM-CNN Hyperparameter Tuning and Training 80:20
Testing	2023-05-23	2023-11-15	LSTM-CNN Price prediction
Prediction	2023-11-16	2023-12-14	Price Prediction 20 days

Table 6. Hyperparameter tuning LSTM-CNN and architecture.

Parameter	Best	Search Space
	Model	
LSTM units	128	min: 32, max: 128, step: 32, sampling: linear
Convolutional filters	32	min: 16, max: 64, step: 16, sampling: linear
Dense units	128	min: 64, max: 256, step: 64, sampling: linear
Learning rate	0.00397	min: 0.0001, max: 0.01, step: No, sampling: log.
Input LSTM leg	Long term data (11): AAPL OHLCV, EMA200, Fama-French 5 Factors	
Input CNN leg	Patterns (15): AAPL Technical indicators, VIX	
Output	20 days prediction	
Target	Close price	
Validation loss	0.003841	

The LSTM-CNN best model (Appendix Figure 14), and previously trained was fed with unseen data (Testing range), the last 20 days' information was removed, and the prediction was executed.

The underlying price is predicted for the next 20 business days, which is effectively 30 calendar days for options. This analysis and prediction will help us to decide the bull/bear/neutral options trading strategy. The prediction resulted in Bullish, with a percentage change of 5.29%, as seen in Figure 9. Compared with the true underlying asset price, both slopes were positive. The main objective is to forecast a 20-day trend, and the model was performing right during this period, the first half was more volatile, while the second half captured a similar pattern in the price evolution.

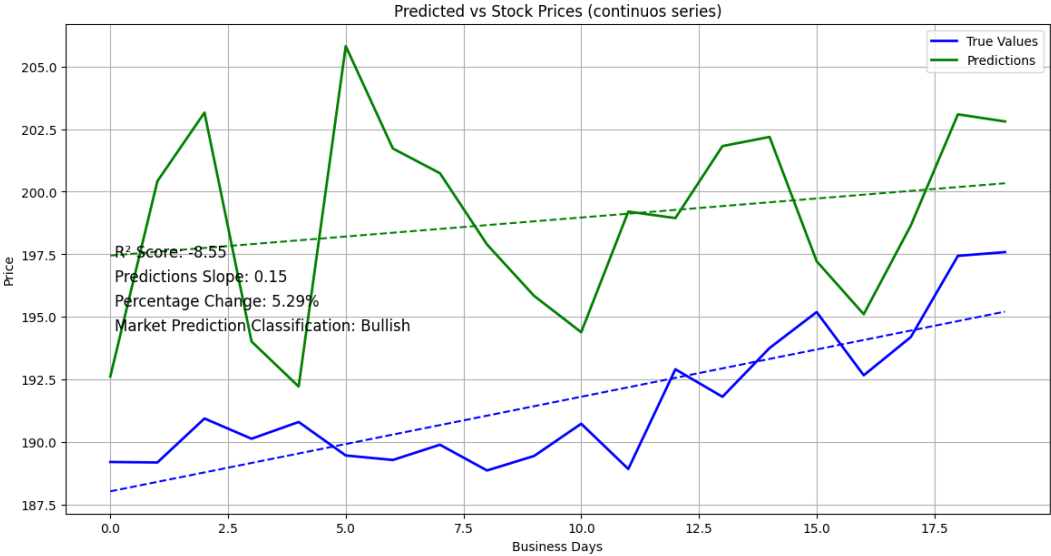


Figure 9. Underlying AAPL Price Prediction (2023-11-16 to 2023-12-14) using LSTM-CNN best model.

4.2. Options strategy evaluation

AAPL historical options chain end-of-the-day quote data with monthly expiries are used. All possible combinations of the Bull Put Spread, Bear Call Spread, and Short Strangle for a given combination of the same expiry date and quote date are analyzed. Table 7 describes the test, train, and prediction split of every strategy.

Table 7. Historical options data split.

Data	From	To
Training	2020-09-01	2023-05-22
Testing	2023-05-23	2023-10-31
Prediction	2023-11-01	2023-12-31

The Open AI Gym Reinforcement Learning Environment will select the strategy combinations based on the trend and high probability of expiring the worthless for a range (from 15 days to 45 days) of days to expiration.

The rewards are compared with random step action, based prediction with PPO, and optimized the PPO model for the best reward. Hyperparameters are tuned to get the best-optimized model with the hyperparameters as Table 8. The plot result suggests that the best PPO outperformed the other two approaches in terms of rewards as per Figure 10.

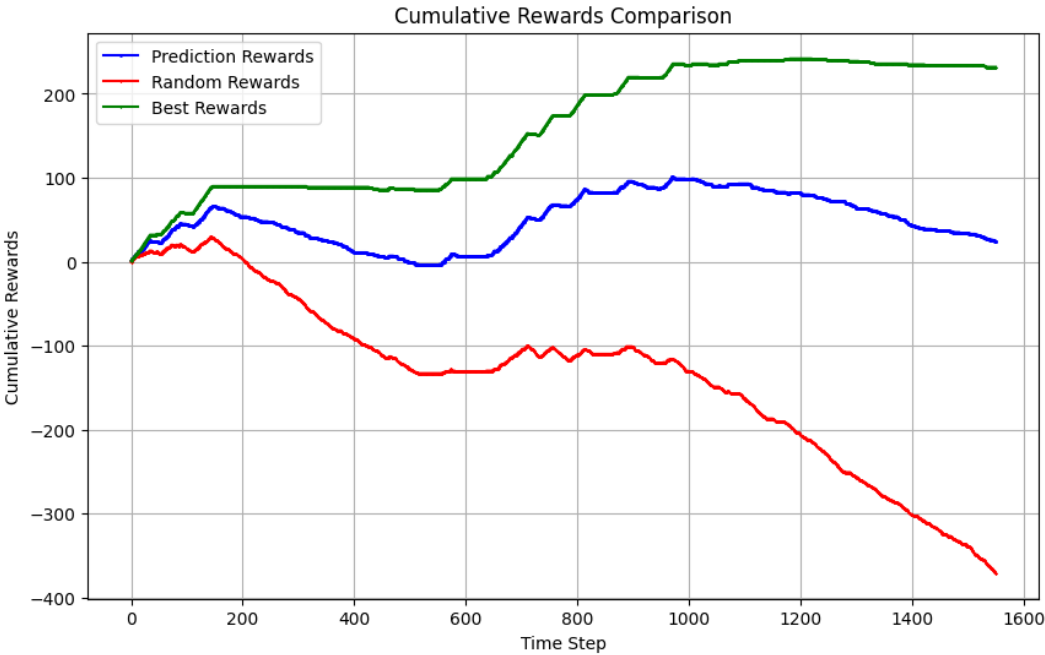


Figure 10. Bull Put Spread Short Strangle Gym Environment Evaluation - Random, Base PPO, and Best PPO.

Table 8. Hyperparameter tuning for PPO Bull Put Spread.

Data	Value
Batch	224
Steps	56
Gamma (Discount Factor)	0.930
Learning Rate	4.273×10^{-5}
Entropy Coefficient	4.48×10^{-6}
Clip Range	0.395
Number of epochs	3
Lambda	0.990
Maximum Gradient Norm	1.089
Value Function Coefficient	0.520

A similar evaluation approach for both Bear Call Spread as mentioned previously for Bull Put Spread is followed. The best PPO model is based on the hyperparameters Table 9 was able to perform well as compared to the other two approaches but could not find many positive rewards as in Figure 11. This could be due to the reason that after the COVID-19 outbreak, the AAPL market was recovering, and there were low instances of bearish markets.

Table 9. Hyperparameter tuning for PPO Bear Call Spread.

Data	Value
Batch	240
Steps	60
Gamma (Discount Factor)	0.880
Learning Rate	1.7×10^{-3}
Entropy Coefficient	9.1×10^{-4}

Clip Range	0.361
Number of epochs	13
Lambda	0.865
Maximum Gradient Norm	2.527
Value Function Coefficient	0.181

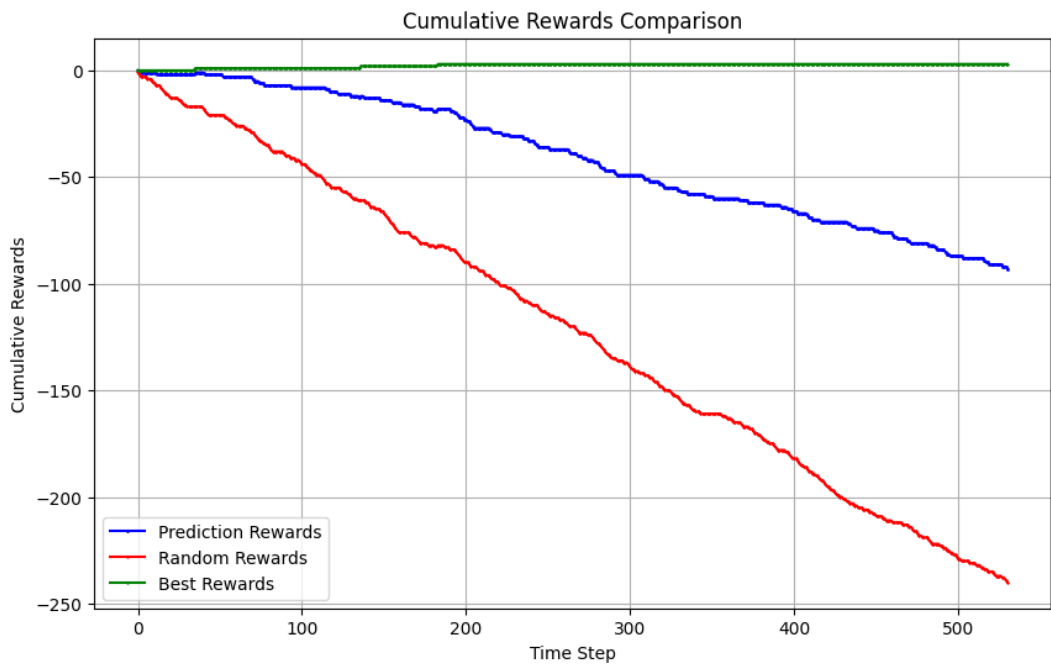


Figure 11. Bear Call Spread Gym Environment Evaluation - Random, Base PPO, and Best PPO.

The strategy evaluation for the Short strangle also follows the similar footsteps to the hyperparameter-tuned approach with parameters as in Table 10 shows higher cumulative rewards as in Figure 12.

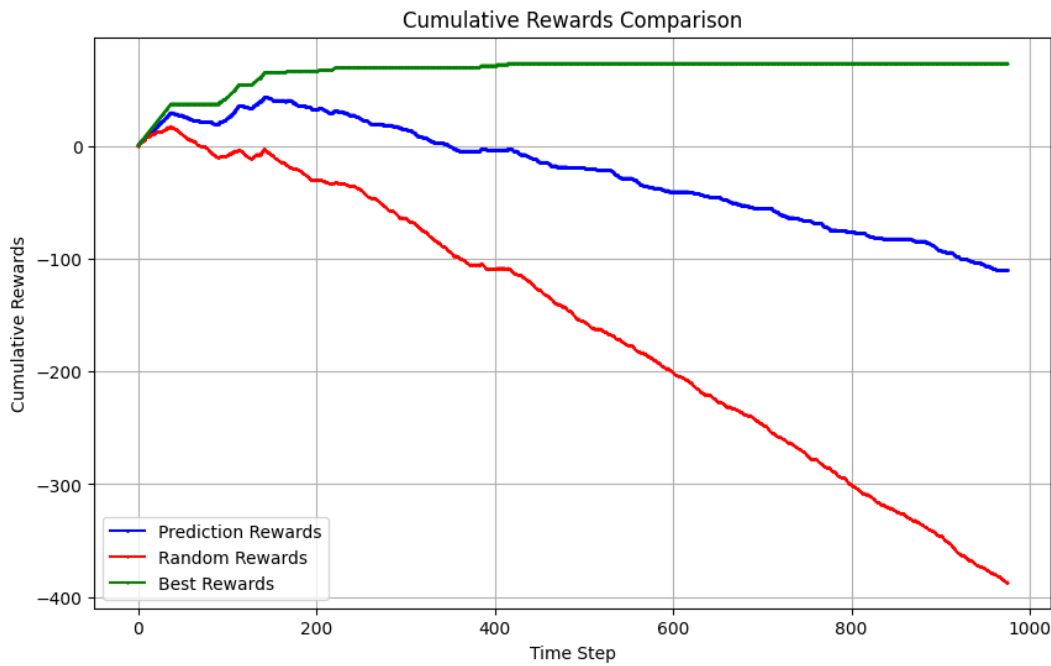


Figure 12. Short Strangle Gym Environment Evaluation - Random, Base PPO, and Best PPO.

Table 10. Hyperparameter tuning for PPO Short Strangle.

Data	Value
Batch	208
Steps	52
Gamma (Discount Factor)	0.893
Learning Rate	1.6×10^{-3}
Entropy Coefficient	5.34×10^{-4}
Clip Range	0.188
Number of epochs	3
Lambda	0.880
Maximum Gradient Norm	2.267
Value Function Coefficient	0.485

4.3. Backtesting

Backtesting is performed on the test data to check how those highly rewarded strategy combinations for each strategy performed on the day of expiry for a single lot. As per the trading summary in Table 11 the Bull Put Spread and Short Strangle have quite a significant number of trades as compared to the Bear call Spread.

Table 11. Backtesting results - Test data.

Strategy	Bull Put Spread		Bear Call Spread		Short Strangle	
	Base	Best	Base	Best	Base	Best
Total Margin (USD)	66678	80825	1023	1023	199330	232930
Total Profit (USD)	6341.9	8062	134	134	12563	15394
Total Trades	209	255	3	3	63	74
Successful Trades	149	181	2	2	51	61
Unsuccessful Trades	60	74	1	1	12	13

4.4. Option Selling with Predicted Trend

The optimized best PPO models are saved for each strategy type and trend to simulate the trading on the unseen data. The hybrid LSTM-CNN predicted the Bullish trend as Figure 9, shows the simulated Bull Put Spread trades for that period, and as per Table 12 shows profit for that short period.

Table 12. Backtesting results.

Strategy	Bull Put Spread
PPO Model	Best/Trade
Total Margin (USD)	1953
Total Profit (USD)	797
Total Trades	7
Successful Trades	7
Unsuccessful Trades	0

5. Discussion

Historical options chain data is not openly available, and this was a challenge which limited exploring other underlying assets that could perform differently under market conditions.

Most reinforcement learning libraries provide support for stocks and crypto trading, which makes it challenging to define and develop a custom options trading RL environment. The deep learning models (hybrid LSTM-CNN) have a random nature that makes the reproducibility of results challenging also. We can extend the test and add more alpha factors to improve the precision in the trend direction.

Future work can use automated trading by getting real-time price and trend predictions from the LSTM-CNN model and adjusting to multiple options positions with the continuous learning of the RL environment.

6. Conclusion

A framework is proposed for options strategies using deep learning methods, the system collects options and stock data, processes data and extracts features, then uses a hybrid LSTM-CNN model to predict the price direction, which can be bullish, neutral, or bearish, then using reinforcement learning, the system selects the best group of strategies (Bear Call Spread, Bull Put Spread or Short Strangle), and backtesting is performed during this forecast period, again using reinforcement learning. This end-to-end development aims to reduce the entry gap for traders and use quantitative analysis to make decisions. The system relies on price direction forecasts, which cannot always be accurate, however, reinforcement learning will make the best possible trades. Also, two of the strategies considered have embedded risk protection (caped risk).

The price trend prediction was aligned with the true price trend during the testing period, and after hyperparameter tuning, this made it possible to increase the profits because the strategy selection was in line with this prediction. Extensive tests need to be conducted, however, the lack of free data, especially the Options Chain makes it difficult, extra computational resources are required to reduce the time for hyperparameter tuning and extend the tests to other assets.

LSTM-CNN prove to be a valuable method to use in financial time series forecasts, and reinforcement learning to decide a trading environment. Future work can improve the factors in price prediction, the LSTM leg is for long-term patterns, and short-term or pattern features in the CNN side, different combinations and additional alphas can improve the model performance to accurately predict the price direction.

The number of possible strategies in options trading makes the use of deep learning methods and quantitative analysis optimal for these problems. Reinforcement learning helps to create an automated trading system without any intervention from humans. Options trading requires a dynamic environment where market conditions change very rapidly. A custom reinforcement learning environment to dynamically capture the high dimensional complexity of the Options Trading with the PPO agent (and MLP policy) was developed.

The backtesting module was developed to assess the options strategies framework. There are very few options trading backtesting frameworks available in the market that can evaluate the multi-leg strategies.

Appendix

The following two figures detail the hybrid LSTM-CNN model and was part of the result hyperparameter tuning process.

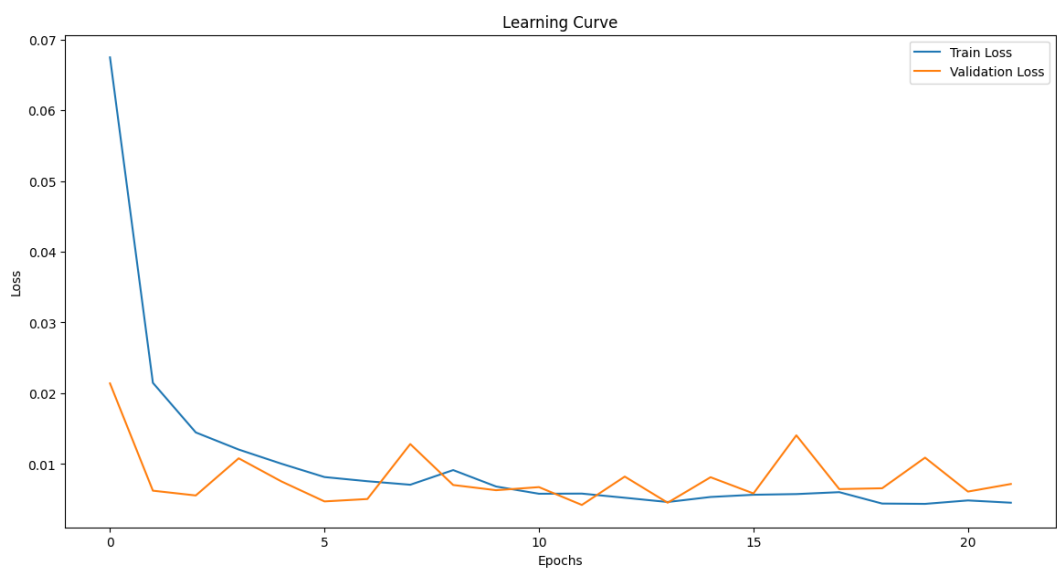


Figure A1. Learning curve for LSTM-CNN model.

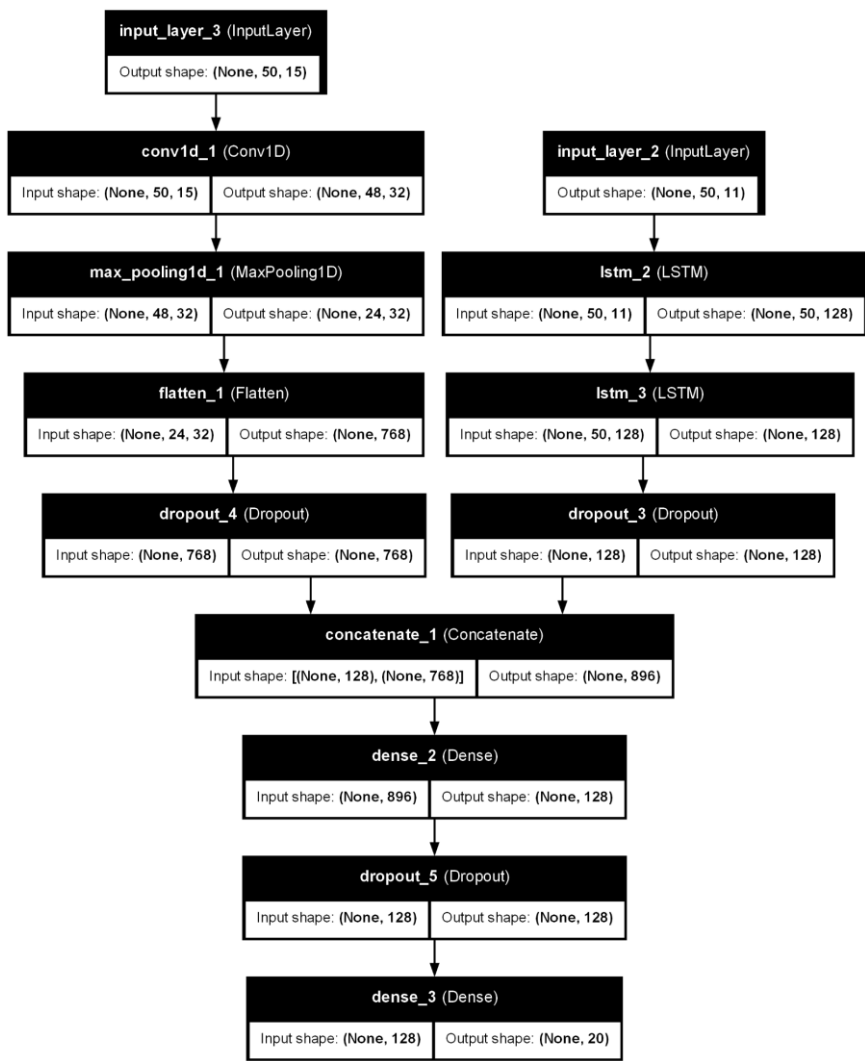


Figure A2. Detailed LSTM-CNN hybrid model.

References

1. P. Ciana, *New Frontiers in Technical Analysis: Effective Tools and Strategies for Trading and Investing*, 1st ed. Bloomberg Press, Sept.2011. [Online]. Available: <https://learning.oreilly.com/library/view/new-frontiers-in/9781576603765/> (accessed Jul. 01, 2024).
2. G. Cohen, *Bible of options strategies, the: the definitive guide for practical trading strategies*, 2nd ed. Pearson, 2005. pp.31,176,180. [Online]. Available: <https://learning.oreilly.com/library/view/the-bible-of/9780133964431/> (accessed Jul. 03, 2024).
3. V. Drakopoulou, "A Review of Fundamental and Technical Stock Analysis Techniques," *Journal of Stock & Forex Trading*, vol. 5, no. 1, Nov. 9, 2016. [Online]. Available: <https://ssrn.com/abstract=3204667>. (accessed Jul. 01, 2024).
4. S. Ravichandiran, *Hands-On Reinforcement Learning with Python*. Packt Publishing, Jun. 2018. [Online]. Available: <https://learning.oreilly.com/library/view/hands-on-reinforcement-learning/9781788836524/> (accessed Jul. 02, 2024).
5. Y. Hilpisch, *Artificial Intelligence in finance*, 1st ed., O'Reilly Online Learning, Oct. 2020. [Online]. Available: <https://learning.oreilly.com/library/view/artificial-intelligence-in/9781492055426/> (accessed Jul. 06, 2024).
6. A. Joshi, B. Venkateswaran, and R. Bhattacharyya, "Options selling using machine learning," *Social Science Research Network*, Apr. 2024, doi: 10.2139/ssrn.4766370. (accessed Jul. 01, 2024).
7. Z. Kakushadze and J. A. Serur, "151 trading strategies", Z. Kakushadze and J.A. Serur. 151 Trading Strategies. Cham, Switzerland: Palgrave Macmillan, an imprint of Springer Nature, 1st Edition (2018), XX, 480 pp. 17-39, 46, 40-60; ISBN 978-3-030-02791-9, Aug. 17, 2018. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3247865 (accessed Jul. 02, 2024).
8. W. Li, "Application of Machine Learning in option Pricing: a review," *Advances in Economics, Business and Management Research/Advances in Economics, Business and Management Research*. Apr. 29, 2022. doi: 10.2991/aebmr.k.220405.035. (accessed Jul. 05, 2024).
9. N. Majidi, M. Shamsi, and F. Marvasti, "Algorithmic trading using continuous action space deep reinforcement learning," *Expert Systems With Applications*, vol. 235, p. 121245, Jan. 2024. [Online]. doi: 10.1016/j.eswa.2023.121245. (accessed Jul. 02, 2024).
10. R. McKeon, "Empirical Patterns of Time Value Decay in Options," *China Finance Review International*, vol. 7, no. 4, pp. 429-449, 2017. Available: <https://www.proquest.com/scholarly-journals/empirical-patterns-time-value-decay-options/docview/2534566154/se-2>. doi: 10.1108/CFRI-09-2016-0108.
11. V. Morris, B. Newman, *A Guide To Investing With Options*, 1st ed. Lightbulb Press, Feb. 2004.
12. G. R. Naufal and A. Wibowo, "Time Series Forecasting Based on Deep Learning CNN-LSTM-GRU Model on Stock Prices," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 71, no. 6, pp. 126-133, Jun. 25, 2023, doi: 10.14445/22315381/IJETT-V71I6P215.
13. Businessline, "Rules to Capture Time Decay," Jul. 24, 2022. [Online]. Available: www.proquest.com/newspapers/rules-capture-time-decay/docview/2693197982/se-2. [Accessed: Jul. 7, 2024].
14. Businessline, "Trading Time Decay in Options," Nov. 6, 2022. [Online]. Available: www.proquest.com/newspapers/trading-time-decay-options/docview/2732145904/se-2. [Accessed: Jul. 7, 2024].
15. X. Wang, J. Li, and J. Li, "A deep learning based numerical PDE method for option pricing," *Computational Economics*, vol. 62, no. 1, pp. 149–164, Jun. 2022, doi: 10.1007/s10614-022-10279-x.
16. W. Wen, Y. Yuan, and J. Yang, "Reinforcement learning for options trading," *Applied Sciences*, vol. 11, no. 23, p. 11208, Nov. 2021, doi: 10.3390/app112311208.
17. M.-E. Wu, J.-H. Syu, and C.-M. Chen, "Kelly-Based Options trading strategies on Settlement Date via supervised learning algorithms," *Computational Economics*, vol. 59, no. 4, pp. 1627–1644, Jan. 2022, doi: 10.1007/s10614-021-10226-2.
18. F. Xu and J. Ma, "Intelligent option portfolio model with perspective of shadow price and risk-free profit," *Financial Innovation*, vol. 9, no. 1, Apr. 2023, doi: 10.1186/s40854-023-00488-0.

19. E. F. Fama and K. R. French, "Production of U.S. Rm-Rf, SMB, and HML in the Fama-French Data Library," Chicago Booth Research Paper No. 23-22, Fama-Miller Working Paper, Dec. 18, 2023. [Online]. Available: <https://ssrn.com/abstract=4629613> or <http://dx.doi.org/10.2139/ssrn.4629613>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.